

1. (currently amended) ~~Method for~~ A method of content level monitoring, auditing, trending, and detection of anomalies in access to information, said information including electronic data on computers, said method of performing an application layer semantic analysis to detect information access anomalies, comprising the steps of:

- a) ~~Capturing of capturing data packets on the network;~~
- b) ~~Filtering filtering the captured data~~ packets to detect meaningful packets representing information content;
- c) ~~Decoding processing~~ packets based on semantics of the ~~an~~ application or protocol;
- d) ~~Analyzing packets to map message information contained in the packet into~~ generating a quantitative representation;
- e) ~~Deriving deriving~~ a content signature from the quantitative representation;
- f) ~~Storing the content, along with the signature and attributes into a database~~
- g) ~~Mining the content database to derive~~ f) deriving a prototypical model of content, users, and time that includes a frequency view of a set of content signatures accessed by a given user, where the set of content signatures are indicative of content that is changing over time; and
- h) ~~Detecting anomalies by finding strong deviations~~ g) detecting an information access anomaly by detecting a given deviation from the prototypical model
- i) ~~Processing anomalies to minimize false alarms and increase the precision of anomalies.~~

2. (currently amended) A ~~The~~ method, according to claim 1, where the ~~filtering is based on protocols and applications of interest~~ the prototypical model also includes a time distribution of a set of content accesses by the given user,

3. (currently amended) A ~~The~~ method, according to claim 2, where protocols include database access protocol such as SQL, file server access protocol such as SMB, application protocols such as smtp, telnet, ftp, rep, http, ldap, J2EE, NET, etc. and

~~applications include Notes, Documentum, Word, etc~~ the prototypical model also includes a location distribution of a set of content accesses by the given user.

4. (currently amended) A ~~The~~ The method, according to claim 1, where the quantitative representation is captured as a content distribution vector that captures a frequency based distribution of key words in the message.

5. (currently amended) A ~~The~~ The method, according to claim 1, where the content signature is computed based on ~~moment statistics such as the n-dimensional~~ a moment statistic,

6. (currently amended) A ~~The~~ The method, according to claim 1, where the content signature is computed as a hash of the information content.

7. (currently amended) A ~~The~~ The method, according to claim 1, where the content signature is computed via a document clustering technique where ~~all~~ documents that ~~classify into one cluster share the same~~ content signatures are classified together.

8. (currently amended) A ~~The~~ The method, according to claim 1, further including storing the information content, the content signature, and one or more attributes, where the attributes include one of: user identity, location of access (~~source and destination IP address~~), time of access, content type (~~e.g. excel document vs. work document~~), content length, content hash, content encoding, and one or more content properties (~~e.g. ownership, time of creating, read/write/execute permissions, encryption, password protection status~~).

9. (currently amended) A ~~The~~ The method, according to claim 1, where mining ~~may be~~ is based on statistical clustering and distance based metrics.

10. (currently amended) A ~~The~~ The method, according to claim 9, where a statistical

~~metrics include~~ metric includes frequency of all content signatures accessed by a user.

11. (currently amended) ~~A~~ The method, according to claim 9, where ~~a~~ statistical ~~metrics include~~ metric includes time of all content signatures accessed by a user.

12. (currently amended) ~~A~~ The method, according to claim 9, where ~~a~~ statistical ~~metrics include~~ metric includes location of all content signatures accessed by a user.

13. (currently amended) ~~A~~ The method, according to claim 1, where ~~the~~ prototypical model is derived by mining a content database ~~may be based on machine learning such as neural networks or rule-based expert systems.~~

14. (currently amended) ~~A~~ The method, according to claim 1, where mining may be augmented by content aging, where information is periodically deleted from the content database.

15. (currently amended) ~~A~~ The method, according to claim 14, where content aging ~~depends on the nature of the~~ is a function of a mining algorithm, ~~the organization, type of information being monitored, users, etc~~ and a type of information being monitored.

16. (currently amended) ~~A~~ The method, according to claim 1, where ~~anomalies are based on combinations of user, content, location, and time entities~~ the information access anomaly is based on a given user accessing given content from a given location at a given time.

17. (currently amended) ~~A~~ The method, according to claim 16, where ~~the~~ information access anomaly is detected by a memory-based deviation where the given content accessed by the given user shows a deviation over ~~the~~ normal content accessed.

18. (currently amended) A ~~The~~ method, according to claim 16, where the information access anomaly is detected by a rare content condition, where a the given user accesses given content that is rarely accessed by the given user.
19. (currently amended) A ~~The~~ method, according to claim 16, where the information access anomaly is detected by a time deviation where a the given user accesses the given content at a time different from historical access by the given user.
20. (currently amended) A ~~The~~ method, according to claim 16, where the information access anomaly is detected by a location deviation where a the given user accesses the given content from a location different from historical access by the given user.
21. (currently amended) A ~~The~~ method, according to claim 1, ~~where anomaly processing includes positive correlation with past security violation events further including processing the information access anomaly.~~
22. (currently amended) A ~~The~~ method, according to claim 4 ~~21~~, where processing the information access anomaly processing includes one of: positive correlation with at least one past security violation event, and negative correlation with a past false alarms or non-events alarm or non-event.
23. (currently amended) A ~~The~~ method, according to claim 1, where a set of consistent anomalies are classified into a pattern of misuse.
24. (currently amended) A ~~The~~ method, according to claim 1, where ~~anomalies can be~~ the information access anomaly is detected in real-time.
25. (currently amended) A ~~The~~ method, according to claim 1, where

information access anomaly detection is used for real-time protection of information.

26. (currently amended) ~~A~~ The method, according to claim 25, where real-time anomaly detection is used for protection via real-time alerts.

27. (currently amended) ~~A~~ The method, according to claim 25, where real-time anomaly detection is used for real-time protection via denial of access.

28. (currently amended) ~~A~~ The method, according to claim 25, where real-time anomaly detection is used for real-time protection via additional user validation.

29. (currently amended) ~~A~~ The method ~~for correlating content, users, time, and space at the information level, developing trends based on information access, and detecting anomalies of information access from confidential information repositories without requiring to know the specific type of information being accessed~~ as described in claim 1, where the data packets are associated with access to a confidential information repository.

30-40 (cancelled)

41. (currently amended) ~~An apparatus for monitoring, trending, and detection of anomalies in access to information, said critical information including electronic data on computers, comprises: a network-based computing device that is used to capture packets, filter data content, decode packets based on protocol and application, derive content signatures, generate historical trends, detect anomalies, and provide real-time access control for performing the method of claim 1.~~

42. (currently amended) ~~An~~ The apparatus of claim 41, where it is implemented on a computing device and connected on a network as a passive tap.

43. (currently amended) ~~An~~ The apparatus claim 41, ~~which is implemented as a network appliance that can derive~~ derives information transparently ~~without requiring logs.~~

44. (currently amended) ~~An~~ The apparatus of claim 41, ~~where it is implemented on an end-user computing device such as a laptop or PC.~~

45. (currently amended) ~~An~~ The apparatus claim 41, ~~where it is implemented as a shim on an application server.~~

46. (currently amended) ~~An~~ The apparatus of claim 41, ~~where it is connected to systems monitoring consoles and user identity systems~~ configured as a computer-readable storage medium having processor-executable instructions encoded thereon.

47. (currently amended) ~~An~~ The apparatus of claim 41, ~~where it is connected to firewalls and other~~ an access control systems system to enable real-time ~~access control for monitoring of~~ anomalous information access.

48. (currently amended) ~~An~~ The apparatus claim 41, ~~which is configured for to~~ implement one or more compliance policies ~~using a simple language.~~

Please add the following new claims:

49. (new) A computer-implemented method of detecting an information access anomaly, comprising:

monitoring data packets indicative of changing content over time;

generating a prototypical model; and

performing a semantic analysis against the prototypical model to identify an application level information access anomaly.

50. (new) A computer program product comprising a computer-readable storage medium encoded with processor-executable program instructions for implementing the method of claim 49.

51. (new) Apparatus including a processor and a computer-readable storage medium, the computer-readable storage medium encoded with processor-executable program instructions for implementing the method of claim 49.